RAW SEQUENCE LISTING   DATE: 09/25/2001
PATENT APPLICATION: US/09/841,758   TIME: 16:11:31

Input Set : N:\Crf3\RULE60\09841758.txt
Output Set: N:\CRF3\09252001\I841758.raw

SEQUENCE LISTING

C-->	3	(1) GENERAL INFORMATION:
	5	(i) APPLICANT: Bandman, Olga
	6	Hawkins, Phillip R.
C-->	8	(ii) TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN
	10	(iii) NUMBER OF SEQUENCES: 5
	12	(iv) CORRESPONDENCE ADDRESS:
	13	(A) ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
	14	(B) STREET: 3174 Porter Drive
	15	(C) CITY: Palo Alto
	16	(D) STATE: CA
	17	(E) COUNTRY: US
	18	(F) ZIP: 94304
	20	(v) COMPUTER READABLE FORM:
	21	(A) MEDIUM TYPE: Diskette
	22	(B) COMPUTER: IBM Compatible
	23	(C) OPERATING SYSTEM: DOS
	24	(D) SOFTWARE: FastSEQ Version 1.5
	26	(vi) CURRENT APPLICATION DATA:
C-->	27	(A) APPLICATION NUMBER: US/09/841,758
C-->	28	(B) FILING DATE: 24-Apr-2001
	30	(vii) PRIOR APPLICATION DATA:
	31	(A) APPLICATION NUMBER: 09/088,641
	32	(B) FILING DATE: 1998-06-02
	34	(viii) ATTORNEY/AGENT INFORMATION:
	35	(A) NAME: Billings, Lucy J.
	36	(B) REGISTRATION NUMBER: 36,749
	37	(C) REFERENCE/DOCKET NUMBER: PF-0163 US
	39	(ix) TELECOMMUNICATION INFORMATION:
	40	(A) TELEPHONE: 415-855-0555
	41	(B) TELEFAX: 415-845-4166
	43	(2) INFORMATION FOR SEQ ID NO: 1:
	45	(i) SEQUENCE CHARACTERISTICS:
	46	(A) LENGTH: 472 amino acids
	47	(B) TYPE: amino acid
	48	(C) STRANDEDNESS: single
	49	(D) TOPOLOGY: linear
W-->	51	(ii) MOLECULE TYPE: None
	53	(vii) IMMEDIATE SOURCE:
	54	(A) LIBRARY:
	55	(B) CLONE: 989953
	57	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
	59	Met Ala Thr Lys Cys Gly Asn Cys Gly Pro Gly Tyr Ser Thr Pro Leu
	60	1 5 10 15
	61	Glu Ala Met Lys Gly Pro Arg Glu Glu Ile Val Tyr Leu Pro Cys Ile
	62	20 25 30
	63	Tyr Arg Asn Thr Gly Thr Glu Ala Pro Asp Tyr Leu Ala Thr Val Asp

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/841,758

DATE: 09/25/2001
TIME: 16:11:31

Input Set : N:\Crf3\RULE60\09841758.txt
Output Set: N:\CRF3\09252001\I841758.raw

```
64            35                    40                   45
65  Val Asp Pro Lys Ser Pro Gln Tyr Cys Gln Val Ile His Arg Leu Pro
66      50                    55                60
67  Met Pro Asn Leu Lys Asp Glu Leu His His Ser Gly Trp Asn Thr Cys
68  65                    70                75                    80
69  Ser Ser Cys Phe Gly Asp Ser Thr Lys Ser Arg Thr Lys Leu Val Leu
70                    85                90                    95
71  Pro Ser Leu Ile Ser Ser Arg Ile Tyr Val Val Asp Val Gly Ser Glu
72              100                   105                  110
73  Pro Arg Ala Pro Lys Leu His Lys Val Ile Glu Pro Lys Asp Ile His
74          115                   120                   125
75  Ala Lys Cys Glu Leu Ala Phe Leu His Thr Ser His Cys Leu Ala Ser
76      130                   135                   140
77  Gly Glu Val Met Ile Ser Ser Leu Gly Asp Val Lys Gly Asn Gly Lys
78  145                   150                   155                   160
79  Gly Gly Phe Val Leu Leu Asp Gly Glu Thr Phe Glu Val Lys Gly Thr
80                    165                   170                   175
81  Trp Glu Arg Pro Gly Gly Ala Ala Pro Leu Gly Tyr Asp Phe Trp Tyr
82              180                   185                   190
83  Gln Pro Arg His Asn Val Met Ile Ser Thr Glu Trp Ala Ala Pro Asn
84          195                   200                   205
85  Val Leu Arg Asp Gly Phe Asn Pro Ala Asp Val Glu Ala Gly Leu Tyr
86      210                   215                   220
87  Gly Ser His Leu Tyr Val Trp Asp Trp Gln Arg His Glu Ile Val Gln
88  225                   230                   235                   240
89  Thr Leu Ser Leu Lys Asp Gly Leu Ile Pro Leu Glu Ile Arg Phe Leu
90                    245                   250                   255
91  His Asn Pro Asp Ala Ala Gln Gly Phe Val Gly Cys Ala Leu Ser Ser
92              260                   265                   270
93  Thr Ile Gln Arg Phe Tyr Lys Asn Glu Gly Gly Thr Trp Ser Val Glu
94          275                   280                   285
95  Lys Val Ile Gln Val Pro Pro Lys Lys Val Lys Gly Trp Leu Leu Pro
96      290                   295                   300
97  Glu Met Pro Gly Leu Ile Thr Asp Ile Leu Leu Ser Leu Asp Asp Arg
98  305                   310                   315                   320
99  Phe Leu Tyr Phe Ser Asn Trp Leu His Gly Asp Leu Arg Gln Tyr Asp
100                   325                   330                   335
101  Ile Ser Asp Pro Gln Arg Pro Arg Leu Thr Gly Gln Leu Phe Leu Gly
102          340                   345                   350
103  Gly Ser Ile Val Lys Gly Gly Pro Val Gln Val Leu Glu Asp Glu Glu
104          355                   360                   365
105  Leu Lys Ser Gln Pro Glu Pro Leu Val Val Lys Gly Lys Arg Val Ala
106      370                   375                   380
107  Gly Gly Pro Gln Met Ile Gln Leu Ser Leu Asp Gly Lys Arg Leu Tyr
108  385                   390                   395                   400
109  Ile Thr Thr Ser Leu Tyr Ser Ala Trp Asp Lys Gln Phe Tyr Pro Asp
110                   405                   410                   415
111  Leu Ile Arg Glu Gly Ser Val Met Leu Gln Val Asp Val Asp Thr Val
112          420                   425                   430
```

RAW SEQUENCE LISTING   DATE: 09/25/2001
PATENT APPLICATION: US/09/841,758   TIME: 16:11:31

Input Set : N:\Crf3\RULE60\09841758.txt
Output Set: N:\CRF3\09252001\I841758.raw

```
113  Lys Gly Gly Leu Lys Leu Asn Pro Asn Phe Leu Val Asp Phe Gly Lys
114          435                 440                 445
115  Glu Pro Leu Gly Pro Ala Leu Ala His Glu Leu Arg Tyr Pro Gly Gly
116      450                 455                 460
117  Asp Cys Ser Ser Asp Ile Trp Ile
118  465                 470
121  (2) INFORMATION FOR SEQ ID NO: 2:
123       (i) SEQUENCE CHARACTERISTICS:
124          (A) LENGTH: 1711 base pairs
125          (B) TYPE: nucleic acid
126          (C) STRANDEDNESS: single
127          (D) TOPOLOGY: linear
W-->  129     (ii) MOLECULE TYPE: None
131      (vii) IMMEDIATE SOURCE:
132          (A) LIBRARY:
133          (B) CLONE: 989953
135      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
137  GCGGCCTCTG ACACCAGCAC AGCAAACCCG CCGGGATCAA AGTGTACCAG TCGGCAGCAT      60
138  GGCTACGAAA TGTGGGAATT GTGGACCCGG CTACTCCACC CCTCTGGAGG CCATGAAAGG     120
139  ACCCAGGGAA GAGATCGTCT ACCTGCCCTG CATTTACCGA AACACAGGCA CTGAGGCCCC     180
140  AGATTATCTG GCCACTGTGG ATGTTGACCC CAAGTCTCCC CAGTATTGCC AGGTCATCCA     240
141  CCGGCTGCCC ATGCCCAACC TGAAGGACGA GCTGCATCAC TCAGGATGGA ACACCTGCAG     300
142  CAGCTGCTTC GGTGATAGCA CCAAGTCGCG CACCAAGCTG GTGCTGCCCA GTCTCATCTC     360
143  CTCTCGCATC TATGTGGTGG ACGTGGGCTC TGAGCCCCGG GCCCCAAAGC TGCACAAGGT     420
144  CATTGAGCCC AAGGACATCC ATGCCAAGTG CGAACTGGCC TTTCTCCACA CCAGCCACTG     480
145  CCTGGCCAGC GGGGAAGTGA TGATCAGCTC CCTGGGAGAC GTCAAGGGCA ATGGCAAAGG     540
146  GGGTTTTGTG CTGCTGGATG GGGAGACGTT CGAGGTGAAG GGGACATGGG AGAGACCTGG     600
147  GGGTGCTGCA CCGTTGGGCT ATGACTTCTG GTACCAGCCT CGACACAATG TCATGATCAG     660
148  CACTGAGTGG GCAGCTCCCA ATGTCTTACG AGATGGCTTC AACCCCGCTG ATGTGGAGGC     720
149  TGGACTGTAC GGGAGCCACT TATATGTATG GGACTGGCAG CGCCATGAGA TTGTGCAGAC     780
150  CCTGTCTCTA AAAGATGGGC TTATTCCCTT GGAGATCCGC TTCCTGCACA ACCCAGACGC     840
151  TGCCCAAGGC TTTGTGGGCT GCGCACTCAG CTCCACCATC CAGCGCTTCT ACAAGAACGA     900
152  GGGAGGTACA TGGTCAGTGG AGAAGGTGAT CCAGGTGCCC CCCAAGAAAG TGAAGGGCTG     960
153  GCTGCTGCCC GAAATGCCAG GCCTGATCAC CGACATCCTG CTCTCCCTGG ACGACCGCTT    1020
154  CCTCTACTTC AGCAACTGGC TGCATGGGGA CCTGAGGCAG TATGACATCT CTGACCCACA    1080
155  GAGACCCCGC CTCACAGGAC AGCTCTTCCT CGGAGGCAGC ATTGTTAAGG GAGGCCCTGT    1140
156  GCAAGTGCTG GAGGACGAGG AACTAAAGTC CCAGCCAGAG CCCCTAGTGG TCAAGGGAAA    1200
157  ACGGGTGGCT GGAGGCCCTC AGATGATCCA GCTCAGCCTG GATGGGAAGC GCCTCTACAT    1260
158  CACCACGTCG CTGTACAGTG CCTGGGACAA GCAGTTTTAC CCTGATCTCA TCAGGGAAGG    1320
159  CTCTGTGATG CTGCAGGTTG ATGTAGACAC AGTAAAAGGA GGGCTGAAGT TGAACCCCAA    1380
160  CTTCCTGGTG GACTTCGGGA AGGAGCCCCT TGGCCCAGCC CTTGCCCATG AGCTCCGCTA    1440
161  CCCTGGGGGC GATTGTAGCT CTGACATCTG GATTTGAACT CCACCCTCAT CACCCACACT    1500
162  CCCTATTTTG GGCCCTCACT TCCTTGGGGA CCTGGCTTCA TTCTGCTCTC TCTTGGCACC    1560
163  CGACCCTTGG CAGCATGTAC CACACAGCCA AGCTGAGACT GTGGCAATGT GTTGAGTCAT    1620
164  ATACATTTAC TGACCACTGT TGCTTGTTGC TCACTGTGCT GCTTTTCCAT GAGCTCTTGG    1680
165  AGGCACCAAG AAATAAACTC GTAACCCTGT C                                   1711
167  (2) INFORMATION FOR SEQ ID NO: 3:
169       (i) SEQUENCE CHARACTERISTICS:
170          (A) LENGTH: 472 amino acids
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/841,758

DATE: 09/25/2001
TIME: 16:11:31

Input Set : N:\Crf3\RULE60\09841758.txt
Output Set: N:\CRF3\09252001\I841758.raw

```
171              (B) TYPE: amino acid
172              (C) STRANDEDNESS: single
173              (D) TOPOLOGY: linear
W--> 175     (ii) MOLECULE TYPE: None
177     (vii) IMMEDIATE SOURCE:
178              (A) LIBRARY: GenBank
179              (B) CLONE: 1374792
181     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
183     Met Ala Thr Lys Cys Gly Asn Cys Gly Pro Gly Tyr Ser Thr Pro Leu
184     1               5                   10                  15
185     Glu Ala Met Lys Gly Pro Arg Glu Glu Ile Val Tyr Leu Pro Cys Ile
186                 20                  25                  30
187     Tyr Arg Asn Thr Gly Thr Glu Ala Pro Asp Tyr Leu Ala Thr Val Asp
188                 35                  40                  45
189     Val Asp Pro Lys Ser Pro Gln Tyr Cys Gln Val Ile His Arg Leu Pro
190         50                  55                  60
191     Met Pro Asn Leu Lys Asp Glu Leu His His Ser Gly Trp Asn Thr Tyr
192     65                  70                  75                  80
193     Ser Ser Cys Phe Gly Asp Ser Thr Lys Ser Arg Asn Lys Leu Val Leu
194                 85                  90                  95
195     Pro Ser Leu Ile Ser Ser Arg Ile Tyr Val Val Asp Val Gly Ser Glu
196                 100                 105                 110
197     Pro Gly Pro Gln Lys Leu His Lys Val Ile Glu Pro Lys Asp Ile His
198                 115                 120                 125
199     Ala Lys Cys Glu Leu Ala Cys Leu His Thr Ser His Cys Leu Ala Ser
200         130                 135                 140
201     Gly Glu Val Met Ile Ser Ser Leu Gly Asp Val Lys Gly Asn Gly Lys
202     145                 150                 155                 160
203     Gly Gly Phe Val Leu Leu Asp Gly Glu Thr Phe Glu Val Lys Gly Thr
204                 165                 170                 175
205     Trp Glu Arg Pro Gly Gly Ala Ala Pro Leu Gly Tyr Asp Phe Trp Tyr
206                 180                 185                 190
207     Gln Pro Arg His Asn Val Met Ile Ser Thr Glu Trp Ala Ala Pro Asn
208                 195                 200                 205
209     Val Leu Arg Asp Gly Phe Asn Pro Ala Asp Val Glu Ala Gly Leu Tyr
210         210                 215                 220
211     Gly Ser His Leu Tyr Val Trp Asp Trp Gln Arg His Glu Ile Val Gln
212     225                 230                 235                 240
213     Thr Leu Ser Leu Lys Asp Gly Leu Ile Pro Leu Glu Ile Arg Phe Leu
214                 245                 250                 255
215     His Asn Pro Ser Ala Thr Gln Gly Phe Val Gly Cys Ala Ser Ala Pro
216                 260                 265                 270
217     Asn Ile Gln Arg Phe Tyr Lys Thr Arg Glu Gly Thr Trp Ser Val Glu
218                 275                 280                 285
219     Lys Val Ile Gln Val Pro Pro Lys Lys Val Lys Gly Trp Leu Leu Pro
220         290                 295                 300
221     Gly Val Pro Gly Leu Ile Thr Asp Ile Leu Leu Ser Leu Asp Asp Arg
222     305                 310                 315                 320
223     Phe Leu Tyr Phe Ser Asn Trp Leu His Gly Asp Leu Arg Gln Tyr Asp
```

```
224                  325              330              335
225  Ile Ser Asp Pro Gln Arg Pro Arg Leu Thr Gly Gln Leu Phe Leu Gly
226              340              345              350   .
227  Gly Ser Ile Val Lys Gly Gly Pro Val Gln Val Leu Glu Asp Glu Glu
228          355              360              365    .
229  Leu Lys Ser Gln Pro Glu Pro Leu Val Val Lys Gly Lys Arg Val Ala
230      370              375              380
231  Gly Gly Pro Gln Met Ile Gln Leu Ser Leu Asp Gly Lys Arg Leu Tyr
232  385              390              395              400
233  Ile Thr Thr Ser Leu Tyr Ser Ala Trp Glu Lys Gln Phe Tyr Pro Asp
234              405              410              415
235  Leu Ile Arg Glu Gly Ser Val Met Leu Gln Val Asp Val Asp Thr Val
236              420              425              430
237  Lys Gly Gly Leu Lys Leu Asn Pro Asn Cys Leu Val Asp Phe Gly Lys
238      .   435              440              445
239  Glu Pro Leu Gly Pro Ala Leu Ala His Glu Leu Arg Tyr Pro Gly Gly
240      450              455              460
241  Asp Cys Ser Ser Asp Ile Trp Ile
242  465              470
244  (2) INFORMATION FOR SEQ ID NO: 4:
246       (i) SEQUENCE CHARACTERISTICS:
247            (A) LENGTH: 472 amino acids
248            (B) TYPE: amino acid
249            (C) STRANDEDNESS: single
250            (D) TOPOLOGY: linear
252      (ii) MOLECULE TYPE: peptide
254     (vii) IMMEDIATE SOURCE:
255            (A) LIBRARY: GenBank
256            (B) CLONE: 227630
258      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
260  Met Ala Thr Lys Cys Thr Lys Cys Gly Pro Gly Tyr Ser Thr Pro Leu
261  1               5               10              15
262  Glu Ala Met Lys Gly Pro Arg Glu Glu Ile Val Tyr Leu Pro Cys Ile
263          20              25              30
264  Tyr Arg Asn Thr Gly Thr Glu Ala Pro Asp Tyr Leu Ala Thr Val Asp
265          35              40              45
266  Val Asp Pro Lys Ser Pro Gln Tyr Ser Gln Val Ile His Arg Leu Pro
267      50              55              60
268  Met Pro Tyr Leu Lys Asp Glu Leu His His Ser Gly Trp Asn Thr Cys
269  65              70              75              80
270  Ser Ser Cys Phe Gly Asp Ser Thr Lys Ser Arg Asn Lys Leu Ile Leu
271              85              90              95
272  Pro Gly Leu Ile Ser Ser Arg Ile Tyr Val Val Asp Val Gly Ser Glu
273              100             105                 110
274  Pro Arg Ala Pro Lys Leu His Lys Val Ile Glu Ala Ser Glu Ile Gln
275          115             120             125
276  Ala Lys Cys Asn Val Ser Ser Leu His Thr Ser His Cys Leu Ala Ser
277          130             135             140
278  Gly Glu Val Met Val Ser Thr Leu Gly Asp Ile Gln Gly Asn Gly Lys
```

VERIFICATION SUMMARY   DATE: 09/25/2001
PATENT APPLICATION: US/09/841,758   TIME: 16:11:32

Input Set : N:\Crf3\RULE60\09841758.txt
Output Set: N:\CRF3\09252001\I841758.raw

L:3 M:220 C: Keyword misspelled or invalid format, [(1) GENERAL INFORMATION:]
L:8 M:220 C: Keyword misspelled or invalid format, [(ii) TITLE OF INVENTION:]
L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:51 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=1
L:129 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=2
L:175 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=3